University of Alberta

Validation of remotely sensed land surface phenology with leafing and flowering records from a citizen science network

by

Logan Macdonald Purdy

A thesis submitted in partial fulfillment of the requirements for a dual degree

Master of Forestry, University of Alberta, Edmonton, Alberta

and

Master of Science in Conservation and Land Management,
Bangor University, Wales

ABSTRACT

Satellites monitor photosynthetic activity, green-up of vegetation in spring, and senescence of leaves in fall at a global scale. These data are used to estimate global and regional primary productivity of vegetation, and to assess the impact of climate change and climate extreme events on ecosystems. Indices of photosynthetic activity calculated from multispectral remote sensing data may, however, be systematically biased or affected by random errors that arise from cloud cover, effects of snow and water, landcover, and vegetation types. Additionally, the timing of satellite observed green-up and senescence is influenced by the method of calculating these dates. Here, I compare remotely sensed land surface phenology with ground phenology observations to evaluate which indices and remote sensing products provide the most accurate assessments of phenology, and to provide corrections for biased estimates. I rely on data from the Alberta PlantWatch citizen science network, which includes more than 57,000 observations of bloom and leaf-out collected from 1987-2016, to ground-truth EVI and NDVI-based phenology estimates from the AVHRR and MODIS sensors. I evaluate a global product from NASA and a regional product specifically designed for forested landcover. Because the study area covers a wide variety of ecosystems and landcover classes, the analysis is stratified by Alberta's ecoregions and main landcover types. The results indicate that green-up estimates are most accurate and unbiased for deciduous forests. Phenology estimates for mixed and coniferous forests are less accurate and have moderate bias (consistently too early or too late) depending on the sensor and remote sensing product. Cropland and grassland estimates have the most variable results among different ecosystems and remote sensing products. The NASA products tend to detect green-up before canopy leaf-out, while a dataset developed for the forested region of Alberta is more accurate. All remote sensing products significantly underestimate the interannual variability of phenology, which carries the risk of underestimating the impact of climate change on ecosystems. However, both bias in the overall timing of phenology and bias in interannual variability can be corrected if ground phenology data are available. For the main ecoregions and landcover types of Alberta, this thesis provides the relevant statistics for bias correction to accurately track phenology changes in response to climate variability and long term climate trends.

ACKNOWLEDGEMENTS

This project would not have been possible without the help and support of many people, to whom I'd like to express my sincerest gratitude.

First, I would like to thank my supervisor Dr. Andreas Hamann. Thank you for assisting with the design of this project, taking the time to answer my questions, providing me with direction, challenging me solve problems on my own, and helping reframe my research to broader issues. I would also like to thank my co-supervisor from Bangor University, Dr. James Gibbons.

Thank you Dr. Elisabeth Beaubien, director of the PlantWatch network. This project would not have been possible without your dedication to initiating, growing, and maintaining PlantWatch. I appreciate your suggestions, guidance, and willingness to answer my questions throughout this project.

Thanks to the NASA EOSDIS Land Processes Distributed Active Archive Center, and to the authors in the Pickell et al. (2017) study for developing and providing the land surface phenology data used in this study. I would also like to thank the volunteers who collected the PlantWatch phenology data.

My masters studies and this research were supported partially by a NSERC Discovery Grant, the College of Alberta Professional Foresters Graduate Award, and the University of Alberta Study Abroad Award.

The Transatlantic Forestry Master (TRANSFOR-M) program has been a rewarding experience from start to finish. I've had the opportunity to meet new people from all over the world, see new places, experience different cultures, and learn a great deal while doing so. Thank you to the program coordinators for giving me the opportunity to participate in such a unique program. Thanks to my classmates at Bangor University, University of Alberta, and other friends who have helped make my time in the TRANSFOR-M program such an enjoyable and memorable experience.

Finally, I would like to thank my parents. Thank you for guiding me, encouraging me, and inspiring me to take risks and seek out new experiences. Your guidance and encouragement have helped me grow, learn, and succeed to the best of my abilities.

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION

1.1 Land surface phenology: a science for global change

Remote sensing is a valuable tool to monitor photosynthetic activity, green-up of vegetation in spring, and senescence of leaves in fall at continental and global scales (Hanes et al. 2014; Helman 2018; Reed et al. 1994). Land surface phenology uses satellite sensors such as AVHRR (Advanced Very High Resolution Radiometer) and MODIS (Moderate Resolution Imaging Spectroradiometer) to monitor surface reflectance using multi-spectral imagery from the visible red and near-infrared portions of the electromagnetic spectrum (Reed et al. 1994). Leaves reflect very little energy in the visible portion of the spectrum due to absorption of photosynthetically active radiation, whereas most energy in the near-infrared is reflected (Huete et al. 2014). Surface reflectance values are transformed into vegetation indices that are related to these absorptive and reflective properties of photosynthesising leaves (Badeck et al. 2004; Delbart et al. 2015; Helman 2018; Huete et al. 2014; Reed et al. 2003). The most commonly used vegetation indices are the Normalised Difference Vegetation Index (NDVI) and the Enhanced Vegetation Index (EVI) (Helman 2018). NDVI is a simple unitless index between -1 and 1 that functions as a normalised ratio between reflectance in the near-infrared band, and reflectance in the visible red band (Huete et al. 2002; Huete et al. 2014; Tucker 1979). EVI is a newer variation of NDVI created on the same premise, with added factors to minimise the influence of aerosol variations and bare soil, and to avoid saturation over dense vegetation (Huete et al. 2002). These indices are used as an indicator of leaf area, fraction of absorbed photosynthetically active radiation (fAPAR), chlorophyll content, and thus photosynthetic capacity (Huete et al. 2014).

Various data processing methods are used to fit annual curves onto discrete vegetation index values in order to fill gaps and smooth the data (Reed et al. 1994). Smoothed time-series are used to identify the dates of important land surface phenology metrics, such as the start of the growing season in spring (SOS), and the end of the growing season or fall senescence (Helman 2018; Reed et al. 1994; White et al. 1997). Start of growth, also known as green-up, is typically identified by when the vegetation index begins increasing or reaches a certain percentage of full summer greenness (Delbart et al. 2015; Helman 2018; Reed et al. 1994; White et al. 1997). Conversely, fall senescence is when the vegetation index falls below a certain percentage of full summer greenness, or reaches its lower thresholds during the transition from fall to winter (Helman 2018; Reed et al. 1994; White et al. 1997). Another important metric is growing season length, which is the number of days between the start and end of the growing season (Helman 2018; Reed et al. 1994; White et al. 1997).

Recent interest in climate change has reinvigorated the scientific discipline of phenology (Donnelly and Yu 2017; Sparks et al. 2009). The timing of leaf-out and bloom in spring is closely linked to preceding temperatures, making changes in phenology an immediate and easily documented impact of climate change on ecosystems (Rosenzweig et al. 2007; Schwartz et al. 2006). Advances in remote sensing have made land surface phenology a valuable tool to monitor the effects of climate on vegetation activity and growing season length. Most studies have found trends towards earlier spring onset, delayed fall senescence, and a lengthened growing season. However, the rate and direction of change has been variable between studies and ecosystems. Spring has advanced across the circumpolar temperate and boreal forest regions, with one study reporting an average advance of 0.16 days/year from 1980-2014 (Park et al. 2016), and another study reporting a more rapid advance of 0.26 days/year from 2000-2014 (Karkauskaite et al. 2017). Trends for fall senescence have been quite variable in North America. One study found a trend towards delayed senescence from 1982-1999, followed by an advance from 2000-2014 in the boreal and subarctic regions (Park et al. 2016). However, a different study found a steadier trend towards delayed senescence from 1982-2008 in the temperate and boreal forest regions (Jeong et al. 2011). The average growing season length in the northern hemisphere has lengthened at an average rate of roughly 0.25-0.40 days/year since 1980, with significant variability in the rate and direction of change between ecosystems (Jeong et al. 2011; Park et al. 2016). Several studies have reported that the trend for spring advance and autumn delay has slowed or become reversed since 2000 (Garonna et al. 2015; Jeong et al. 2011; Park et al. 2016).

Phenology is not only of interest as an ecosystem response to climate change, but also as a driver and feedback mechanism to global climate cycles. Phenology is an important driver of global cycles with influences on surface albedo, water and energy fluxes, primary production, and carbon cycles (Richardson et al. 2013). Growing season length and plant productivity is closely related to carbon sequestration, making phenology an important factor in the potential of ecosystems to mitigate climate change (Gu et al. 2003; Leinonen and Kramer 2002; White et al. 1999). Primary production, the rate at which plants capture and store carbon dioxide from the atmosphere (Xiangming et al. 2014), is related to absorbed solar energy (Running et al. 2004). Given that vegetation indices such as NDVI and EVI are proxies of absorbed solar energy, land surface phenology can be used to estimate plant productivity and study the feedbacks between plants, phenology, and the earth's climate system (Running et al. 2004). The annual and seasonal changes in NDVI have been shown to correspond with annual and seasonal changes in gross and net primary production from other data sources (Goward et al. 1985; Schloss et al. 2001). Matching the trend of warming, several studies have demonstrated that the peak and amplitude of NDVI curves increased from 1980-2000, suggesting a substantial increase in plant growth and photosynthetic activity (Myneni et al. 1997; Tucker et al. 2001; Zhou et al. 2001). Park et al. (2016) used NDVI-based estimates to demonstrate that 42% of the circumpolar boreal region experienced a significant trend towards increased productivity from 1980-2014, which increased total gross primary productivity by 20%. Like trends for start and end of season, the rate of increase in primary production slowed in the latter period of this study (2000-2014) (Park et al. 2016).

While land surface phenology is an effective tool to monitor vegetation and related processes at continental and global scales, there are several sources of error that affect the precision and accuracy of land surface phenology estimates. Monitoring green-up can be complicated when the satellite pixel contains a mosaic of landcover types, non-vegetated areas, anthropogenic landcover types such as cropland and urban areas, or waterbodies (Delbart et al. 2015). Cloud cover creates temporal and spatial gaps in data (White et al. 2014; Zhang et al. 2006), and spring snowmelt exposes bare soil and understorey leaf litter which can increase the vegetation index, creating a false impression of greening-up (Fisher and Mustard 2007; Reed et al. 1994; White et al. 2009). Different vegetation indices, curve smoothing methods, and metric retrieval methods have been shown to produce SOS dates that vary by up to two months for the same pixels (Helman 2018; Misra et al. 2016; White et al. 2009). These factors cause uncertainty on the accuracy of land surface phenology metrics, and associated estimates of growing season length and primary productivity.

1.2 Ground phenology: an opportunity to ground truth remote sensing

An important aspect of land surface phenology research is to establish relationships between the annual changes in vegetation indices and the phenological events driving these changes (Badeck et al. 2004). Ecosystem carbon fluxes and productivity are closely linked to specific plant phenophases, such as the unfolding and maturation of leaves in spring, and the senescence of leaves in fall (D'Odorico et al. 2015; White et al. 1999). Determining how well land surface phenology predicts the timing of these phenophases is crucial to accurately assess these associated ecosystem processes. Thus, it is useful to compare and calibrate land surface phenology data with leafing and flowering data in order to derive bias correction for different landcover and vegetation types, and to select the vegetation index and SOS retrieval method that provides the most accurate assessment of phenology for a study region.

Prior to the use of satellites, phenology was traditionally documented using in-situ observations of individual plants, with a focus on specific phenophases such as flowering and leafing (Badeck et al. 2004; Cleland et al. 2007). The scale and quantity of data required to adequately monitor ground phenology is beyond the capability of scientists alone (Feldman et al. 2018), which has led to the creation of phenology networks that rely on citizen scientists for data collection. These citizen science networks include the PlantWatch network in Canada (Beaubien and Hamann 2011a), the USA National Phenology Network (Betancourt et al. 2007), and many in Europe such as Nature's Calendar in the UK (Collinson and Sparks 2008). Ground phenology data are limited by inconsistencies in observer protocols, temporal and spatial gaps in observations, and the limited spatial distribution of potential observers (Cleland et al. 2007; Kross et al. 2011; Misra et al. 2016).

Compared to ground phenology, land surface phenology is more effective at providing temporally and spatially contiguous records of phenology, and is particularly useful to provide records for areas where there are few or no potential observers (Badeck et al. 2004; Cleland et al. 2007; Kross et al. 2011). However, it is important to ensure that land surface phenology metrics reflect the annual patterns of ground phenology, and studies comparing the two have had variable results. The method to extract land surface phenology metrics can have a considerable impact on the relationship between these metrics and ground phenology. Some methods have been shown to result in strong correlations with little bias or error, while other methods have weak correlations and significant bias (Schwartz et al. 2002; White et al. 2014; White et al. 2009). Other studies have found that leaf-out observations can closely match land surface phenology green-up dates when observations are from forested pixels, while observations from partially forested or unforested pixels do not correlate well (Delbart et al. 2015; Pouliot et al. 2011).

Estimates on the rate of spring advance have also been more variable from remote sensing than from ground phenology data. Ground phenology studies have had relatively consistent findings, with most studies reporting that spring is advancing at a rate of 2-5 days per decade in temperate zones across the planet (Ahas et al. 2002; Badeck et al. 2004; Beaubien and Hamann 2011b; Menzel et al. 2006; Root et al. 2003; Schwartz et al. 2006; Schwartz and Reiter 2000). The variability in results from land surface phenology studies is likely due to the range of methods used. For example, different smoothing techniques and metric retrieval methods can result in opposite trends for spring start in the same study region (Buitenwerf et al. 2015; Misra et al. 2016). Additionally, different vegetation indices can result in differing rates and opposite directions of change for spring start in the same ecosystems (Karkauskaite et al. 2017). Different land surface phenology methods are sensitive to the phenology of different species (Misra et al. 2016), and the rate of change in spring phenology is highly variable among species with some changing in the opposite direction of the warming pattern (Fitter and Fitter 2002; Menzel et al. 2006). Thus, distinct land surface phenology methods are sensitive to different ecosystem characteristics, which may have differing responses to climate change.

Discrepancies between land surface phenology and ground phenology arise due to these two techniques monitoring related, albeit qualitatively different processes. While ground phenology has a plant and phenophase specific focus, land surface phenology captures the aggregated green-up of many species within moderate to coarse spatial scales depending on the pixel size (Badeck et al. 2004; Cleland et al. 2007; Helman 2018). Thus, land surface phenology metrics are not a direct indicator of any ground based phenophase, but are rather an indication of a shift in ecosystem dynamics (Reed et al. 1994). Combining both phenology approaches offers the opportunity to better understand how ground phenology influences satellite estimates of green-up.

1.3 Study objectives

This study evaluates the suitability of different land surface phenology products to detect spring start in northern ecosystems by comparing with ground phenology data collected by the Alberta PlantWatch citizen science network from 1987-2016. This is a more northerly region than other similar North American studies (Delbart et al. 2015; Pouliot et al. 2011; White et al. 2014; White et al. 2009), which is important given that climate change is expected to occur more rapidly in northern ecosystems (IPCC 2014). The Alberta PlantWatch network is among the most northerly phenology networks in North America, giving the unique opportunity to evaluate the relationship between land surface phenology and ground phenology for landscape types that other studies have not considered. Vegetation index, computational method to derive the land surface phenology metric, and landcover are all expected to influence the relationship between ground phenology and land surface phenology metrics. Thus, I stratify the study area into ecological zones based on the Alberta Natural Subregion classifications, and test different landcover types separately within each ecoregion. I calculate mean green-up dates from three land surface phenology datasets for selected landcover types in each ecoregion, using 50 selected grid cells of the chosen landcover type. These green-up dates are then compared against the average date of ground phenology events in the selected ecoregions, calculated using a best linear unbiased prediction model for common species in the Alberta PlantWatch database. I evaluate the accuracy and bias of the land surface phenology green-up estimates in order to assess whether they accurately reflect spring start from ground phenology data. I also provide the relevant statistics and equations necessary to correct biased estimates of phenology.

2. METHODS

2.1 Alberta PlantWatch data

The ground phenology data used in this study were collected by the Alberta PlantWatch citizen science network, which was initiated in 1987 to track spring plant phenology in the province of Alberta (Canada). Citizen volunteers were recruited through newspaper articles, government newsletters, public talks, radio interviews, and conference posters (Beaubien and Hamann 2011a). Volunteers observe and report the calendar date of the following phenophases for common and easily identifiable plant species across Alberta: first bloom (first three flowers open in three different places of a woody shrub/tree, or first flowers open in a patch of herbaceous plants), mid bloom (50% of flower buds open), full bloom (90% of flower buds open), leaf-out (first leaves unfurled in 3+ places on the tree/shrub). Observations protocols were adjusted in 2002 to better match those in Europe. Accordingly, leaf-out observations began, and full bloom ceased to be recorded (Beaubien and Hamann 2011a).

2.2 Stratification of phenology data

As of 2016, the Alberta PlantWatch database includes over 57,000 records for 30 species taken by roughly 700 observers. Observations have also been taken in all of Alberta's 21 Natural Subregions (Figure 1a). In order to select ecoregions for this study, I performed an overlay analysis to identify the best sampled Natural Subregions within the PlantWatch data. Most records are clustered in the central portion of the province (Figure 1a), with the Central Parkland and Dry Mixedwood regions accounting for 60% of the total observations. Based on the number of sampling locations and total number of observations in each Natural Subregion, this study was stratified into six ecoregions: Central Parkland (726 locations; 20,846 observations), Dry Mixedwood (478 locations; 13,962 observations), Foothills Parkland (144 locations; 5,134 observations), Grasslands (313 locations; 6,503 observations), Central Mixedwood (165 locations; 2,463 observations), and Montane (258 locations; 3,900 observations). The Grasslands region is a combination of the Foothills Fescue, Dry

Mixedgrass, and Mixedgrass Natural Subregions, all three of which are within the Grasslands Natural Region (Natural Regions Committee 2006). These ecoregions have 52,808 total phenology observations, which is 91.5% of the PlantWatch database.

Rather than correlating individual phenology observations with remote sensing data from individual grid-cells, I calculated average dates of ground observed phenophases for each ecoregion. Instead of taking regional averages of PlantWatch data, a mixed effects model was used to minimize potential issues due to small sample size, observer bias, and outliers. The mixed effects model gives consideration to the collinearity between phenophases (first, mid, and full bloom, and leaf-out for aspen), and between observations of different species to improve the accuracy of the estimated average ground phenology date. A best linear unbiased prediction (BLUP) mixed model was created with the phase as the predictor, and the year, ecoregion, and species as random effects with the ASReml package in R version 3.4.0 (Butler 2009; R Core Team 2013). The model estimated the date of first bloom, mid-bloom, full bloom, and leaf-out (aspen only) for the nine best sampled species in the PlantWatch data for the six selected ecoregions (Table 1). With the removal of lesser sampled species, this subset included 35,117 individual records, 61% of the PlantWatch database. Each of the nine species has more than 70 records from each ecoregion, except for golden bean for which there are only 4 observations in the Central Mixedwood region (Appendix Table A1).

Table 1. Number of Alberta PlantWatch records from the six chosen ecoregions, by phenophase, for the nine common species used in this analysis.

Species		Bloom-phase				
Common name	Scientific name	First	Mid	Full	Leaf-out	Total
Saskatoon	*Amelanchier alnifolia*	2,044	1,818	1,577		5,439
Early blue violet	*Viola adunca*	1,608	1,420	1,291		4,319
Prairie crocus	*Anemone patens*	1,550	1,285	1,112		3,947
Aspen poplar	*Populus tremuloides*	1,320	1,126	819	601	3,866
Chokecherry	*Prunus virginiana*	1,419	1,226	1,009		3,654
Northern bedstraw	*Galium boreale*	1,344	1,219	1,025		3,588
Golden bean	*Thermopsis rhombifolia*	1,357	1,203	1,004		3,564
Yarrow	*Achillea millefolium*	1,370	1,171	948		3,489
Star-flowered Solomon's seal	*Maianthemum stellatum*	1,310	1,089	852		3,251

Like the approach used for the ground phenology data, I calculated the average green-up date for each of the selected ecoregion – landcover combinations. The location of PlantWatch observations are not necessarily in locations spatially representative of the larger landscape, as many are in agriculture land, built-up areas, or near waterbodies. Comparing individual point observations with the green-up date of the corresponding pixel would thus introduce unnecessary bias and error. Instead, I selected grid cells representative of select landcover to extract the green-up dates in each ecoregion, using 250m resolution MODIS landcover data (CEC 2013). Fifty points were created for each chosen landcover type in the six ecoregions (Figure 1b). The locations of the points were systematically chosen to provide an even spatial representation within the ecoregion, and to not be located near boundaries of other landcover types in order to avoid mosaic pixels.

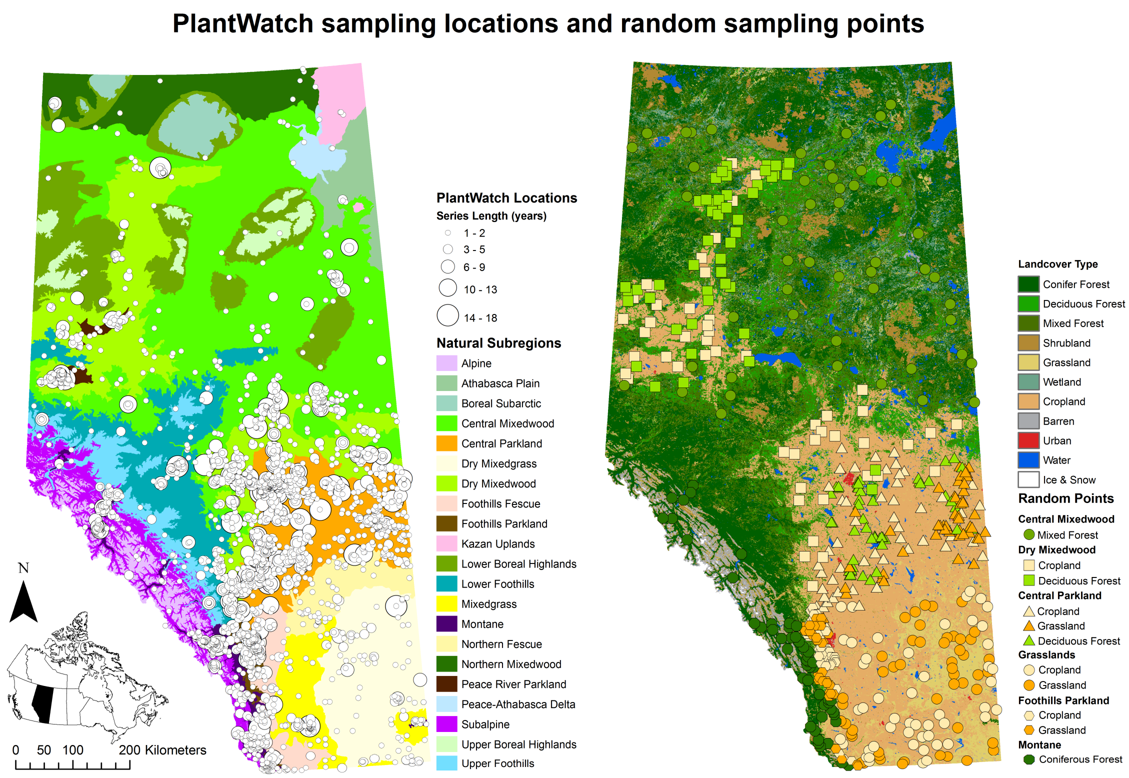

Figure1. (a) PlantWatch sampling locations in Alberta by Natural Subregion (series length is the number of years from 1987-2016 for which phenology observations have been reported at each location); (b) random sampling points created for extracting the land surface phenology green-up dates by ecoregion and landcover type. The Grasslands region in this study is a combination of the Dry Mixedgrass, Mixedgrass, and Foothills Fescue Natural Subregions.

2.3 Land surface phenology data

The NASA MEaSUREs Vegetation Index and Phenology global NDVI (Normalized Difference Vegetation Index) and EVI (Enhanced Vegetation Index) products were used in this study. These data were developed from the NASA Daily Vegetation Index and Phenology product, and include start of season (SOS) estimates at a 5600m pixel size from AVHRR (Advanced Very High Resolution Radiometer; 1981-1991) and MODIS (Moderate Resolution Imaging Spectroradiometer; 2000-2014) sensors (Didan and Barreto 2016a, b;

Didan et al. 2018). These datasets use a modified half-maximum method to detect green-up, where SOS is the day that the vegetation index surpasses a 0.35 ratio between the minimum and maximum vegetation index values for that year (Didan et al. 2018). More detailed data processing information is available in Didan et al. (2018).

A finer resolution (250m pixel size) Alberta-wide NDVI-based land surface phenology product was also used (hereafter referred to as “Pickell NDVI”). This dataset used the half-maximum method (0.5 ratio) to determine SOS from eight-day MODIS imagery composites (Pickell et al. 2017). More detailed data processing information is available in Pickell et al. (2017). The differences in vegetation index, SOS retrieval method, and pixel resolution between these land surface phenology products clearly results in different green-up dates for some locations across Alberta (Figure 2).

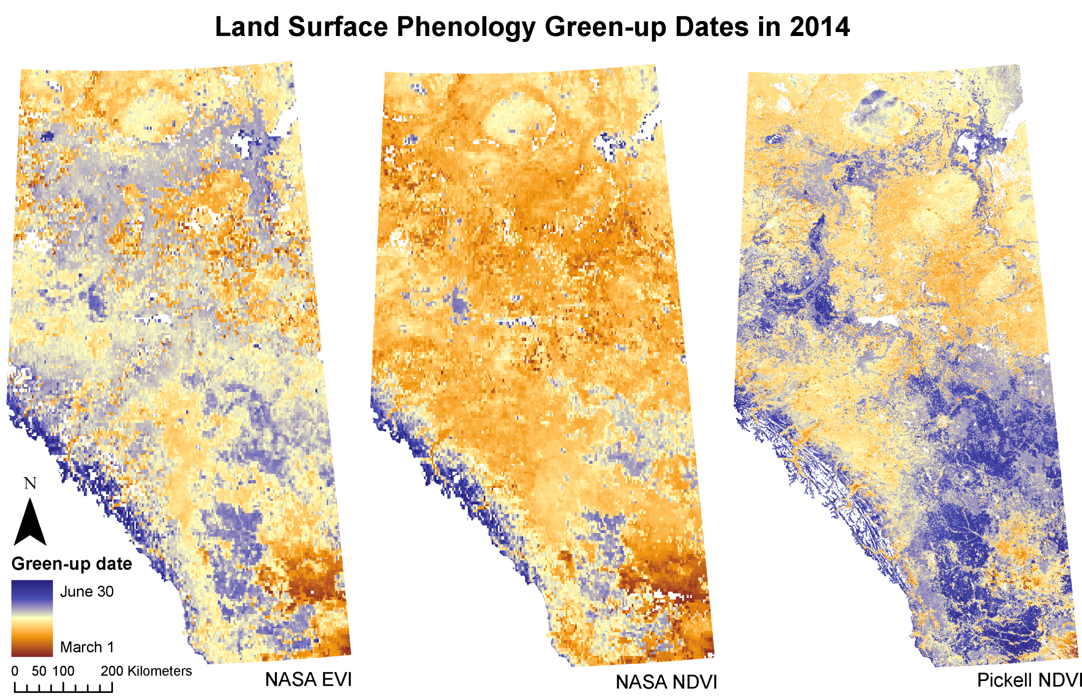

Figure 2. Examples of land surface phenology green-up dates for the year 2014 (Pickell NDVI aggregated to 1km pixels for clarity).

The 550 random points shown in Figure 1b were used to extract the green-up dates of the corresponding pixels for the NASA EVI and NDVI datasets from 1987-2014. The Pickell NDVI green-up dates were extracted from 2000-2016 using these points, plus an additional 125 points created to account for poor data coverage in a few years for certain ecoregions. Green-up dates earlier than day-of-year (doy) 60 were deemed to be erroneous, and were removed from analysis. Other studies have reported doy ~60 as the earliest satellite observed green-up in Alberta (Cui et al. 2017; Pickell et al. 2017). The individual green-up dates were aggregated by landcover type within each ecoregion to determine the mean green-up date for each ecoregion – landcover combination, annually, for each land surface phenology dataset.

2.4 Statistical analysis

The Pearson correlation coefficient (r) was calculated to test the strength of the linear relationship between the land surface phenology mean green-up dates and the estimated ground phenology dates for first bloom of all nine species and leaf-out for aspen. This was done separately for each land surface phenology dataset, in each of ecoregion – landcover combination. The NASA datasets were evaluated separately from 1987-1999 (AVHRR data) and 2000-2014 (MODIS data), and the Pickell NDVI dataset was evaluated from 2000-2016. Due to the large number of possible comparisons, I selected one phenophase to report for each ecoregion. In each case, the chosen phenophase was from a common species with relatively strong correlations with the land surface phenology data, and relatively similar first bloom or leaf-out dates to green-up. I report the correlation coefficient (r) and significance from a one-way positive correlation test for the selected phenophase. The root mean square error (RMSE) was calculated as a measure of bias between the remote sensing green-up dates and ground observed phenology dates. RMSE is a description of the average difference between the green-up estimates and the ground phenology dates (Willmott 1982), and is calculated as follows:

$$(1)\quad \mathrm{RMSE} = \sqrt{\frac{\sum_{i=1}^{N}(S-G)^2}{N}}$$

where S is the satellite observed green-up date, G is the ground observed date for the chosen phenophase, and N is the number of years that the time-series covers. Datasets with green-up estimates similar to the ground phenology dates would have low RMSE, while biased green-up estimates would have high RMSE.

Given that land surface phenology green-up estimates are intended to predict the start of growth and photosynthesis on the ground, I report the average difference in timing between green-up and two phenophases. I calculated the mean lag, which is the mean difference between the green-up date and the ground phenology date. This was calculated relative to aspen first bloom and aspen leaf-out. Aspen is a widespread species throughout Alberta, and is among the best recorded species in the PlantWatch database (Table 1). Aspen bloom is also among the earliest spring phenophases, making it useful as an indicator spring start, and aspen leafing is expected to be an important factor in satellite observed green-up.

3. RESULTS

3.1 Phenology sequence and model performance

The estimated average first bloom date for the nine chosen species (and aspen leaf-out) ranges from late March (doy ~85) to early July (doy 180+) in the Central Parkland region (Figure 3). The earliest spring phenophases are the synchronized bloom of aspen and prairie crocus, after which there is roughly a 20 day lag before aspen leaf-out. There is clearly a degree of correlation across the range of phenophases, though events that occur in a similar time period are more related. Aspen and prairie crocus bloom appear to have greater interannual variability than later phenophases.

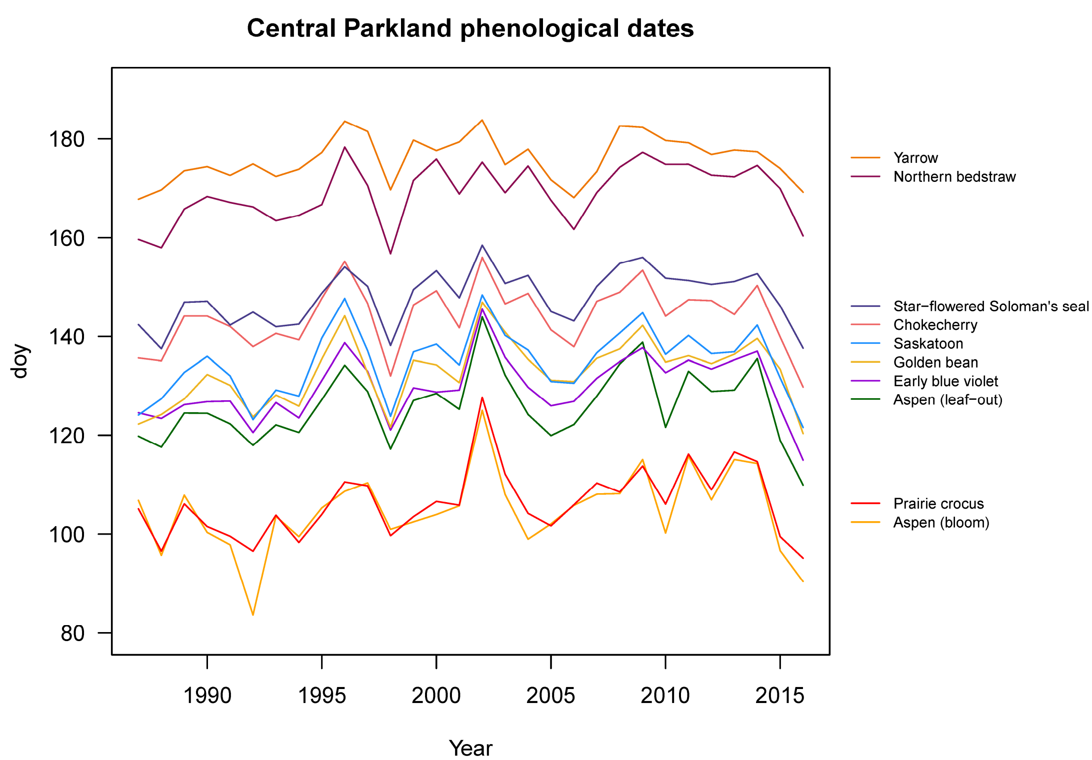

Figure 3. Phenology sequence (doy: day of year) of the first bloom and aspen leaf-out mean date estimate in the Central Parkland region of Alberta.

The mixed effects model produced precise predictions for ground phenology dates, as supported by the low standard errors from the model predictions. Based on the correlation coefficients and the timing of the green-up estimates, aspen leaf-out was chosen as the phenophase to report in the Central Parkland, Dry Mixedwood, and Foothills Parkland regions. Saskatoon first bloom was chosen for the Grasslands region, and aspen first bloom was chosen for the Central Mixedwood and Montane regions. The average standard error for the aspen leaf-out estimate is 2.7 in the intensively sampled Central Parkland region (max: 3.8), 2.8 in the Dry Mixedwood region (max: 3.8), and 3.3 in the Foothills Parkland region (max: 4.0). Saskatoon bloom in the Grasslands region has an average standard error of 1.7, with a maximum of 2.6. Aspen bloom has an average standard error of 2.5 in the Central Mixedwood region (max: 3.7), and 2.4 in the Montane region (max: 2.9). Aspen leaf-out has higher standard errors than first bloom of all species as these records have only been collected since 2002, and due to there being relatively few observations of leaf-out compared to first bloom of other species (Table 1).

3.2 Correlation and RMSE

The strength of the linear relationship between land surface phenology green-up dates and ground phenology varies between land surface phenology products, between ecoregions, and between landcover types within ecoregions (Table 2). Some combinations are strongly correlated with low RMSE, while others have weak or no correlation. Strong correlations are also found across a wide range of other phenophases than those reported, including some that occur much later than green-up (Appendix Table A2).

Table 2. Statistics for the accuracy and precision of the green-up estimate in each ecoregion-landcover combination, for a selected phenophase. The root mean squared error (RMSE) is an alternative measure of bias, and the Pearson correlation coefficient (r) is reported as a measure of precision for the relationship between the remotely sensed sensing green-up and observed phenology (* denotes statistical significance to p<0.05, ** p<0.01).

	MODIS EVI (NASA)		MODIS NDVI (NASA)		MODIS NDVI (Pickell)		AVHRR EVI (NASA)		AVHRR NDVI (NASA)	
Ecoregion/landcover/species	**RMSE**	**r**	**RMSE**	**r**	**RMSE**	**r**	**RMSE**	**r**	**RMSE**	**r**
Central Parkland (aspen leaf-out)										
Deciduous	15.3	0.56*	30.9	0.30	5.8	0.75**	10.5	0.26	20.6	0.37
Cropland	6.2	0.42	21.1	0.32	16.9	0.46*	6.3	0.24	12.2	0.16
Grassland	16.7	0.11	33.2	0.18	13.7	0.37	10.7	0.22	23.0	0.40
Dry Mixedwood (aspen leaf-out)										
Deciduous	10.6	0.64**	28.9	0.50*	4.0	0.86**	5.7	0.73**	15.2	0.29
Cropland	8.2	0.57*	23.1	0.52*	13.7	0.63**	4.2	0.64**	11.0	0.56*
Foothills Parkland (aspen leaf-out)										
Cropland	14.0	0.71**	26.8	0.66**	12.1	0.47*	14.4	0.77**	21.2	0.75**
Grassland	12.3	0.69**	22.3	0.64**	12.5	0.33	12.3	0.81**	19.1	0.89**
Grasslands (Saskatoon bloom)										
Cropland	11.1	0.14	24.1	-0.02	17.1	0.09	10.5	0.68**	20.0	0.66**
Grassland	36.6	0.00	45.4	-0.07	20.7	0.12	33.4	0.61*	39.3	0.64**
Central Mixedwood (aspen bloom)										
Mixed forest	13.1	0.30	16.6	0.16	10.1	0.65**	16.2	0.34	7.1	0.40
Montane (aspen bloom)										
Conifer forest	21.5	0.03	7.0	0.61**	8.5	0.58**	18.4	0.18	10.4	-0.15

The Pickell NDVI green-up dates are strongly correlated with aspen leaf-out for deciduous forests in the Central Parkland and Dry Mixedwood regions, with low RMSE (Table 2). This dataset's green-up dates from cropland landcover are strongly correlated with aspen leaf-out in the Dry Mixedwood region, and have weaker significant correlations for cropland landcover in the Central Parkland and Foothills Parkland regions. The Pickell NDVI green-up dates are also fairly strongly correlated with aspen bloom for mixed and coniferous forests in the Central Mixedwood and Montane regions respectively, with RMSE of 8-10 days.

The NASA MODIS NDVI green-up dates tend to have weaker correlations with ground phenology than Pickell NDVI for forested ecoregions (Table 2). However, the correlation strength is quite similar for these two MODIS NDVI datasets for the coniferous forest landcover type in the Montane region. The NASA MODIS NDVI data are also strongly correlated with aspen leaf-out for the cropland and grassland landcover in the Foothills Parkland region, though with high RMSE. The NASA AVHRR data have moderate correlations for the cropland and grassland landcover types the Grasslands region, and strong correlations for these landcover types in the Foothills Parkland region, also with high RMSE.

The NASA MODIS EVI green-up dates have moderate correlations with aspen leaf-out for deciduous forests in the Central Parkland and Dry Mixedwood regions, with moderate RMSE. Correlations are strong for cropland and grassland landcover in the Foothills Parkland region, also with moderate RMSE. The NASA AVHRR EVI data have similar correlations to the AVHRR NDVI data for cropland and grassland landcover in the Foothills Parkland and Grasslands regions, with lower RMSE. The NASA AVHRR EVI data are also strongly correlated with aspen leaf-out for deciduous forest and cropland landcover in the Dry Mixedwood region, with low RMSE.

3.3 Interannual variability

Comparing time-series of the green-up dates against ground phenology reveals that the NASA land surface phenology products have far less interannual variation (Figure 4). This is even more pronounced for the AVHRR based estimates prior to 2000. Estimates from the Pickell NDVI product appear to better represent the interannual variability of ground phenology in the forested ecoregions. The NASA data follow the trends of the ground phenology in several regions, though early and late years are not nearly as distinct.

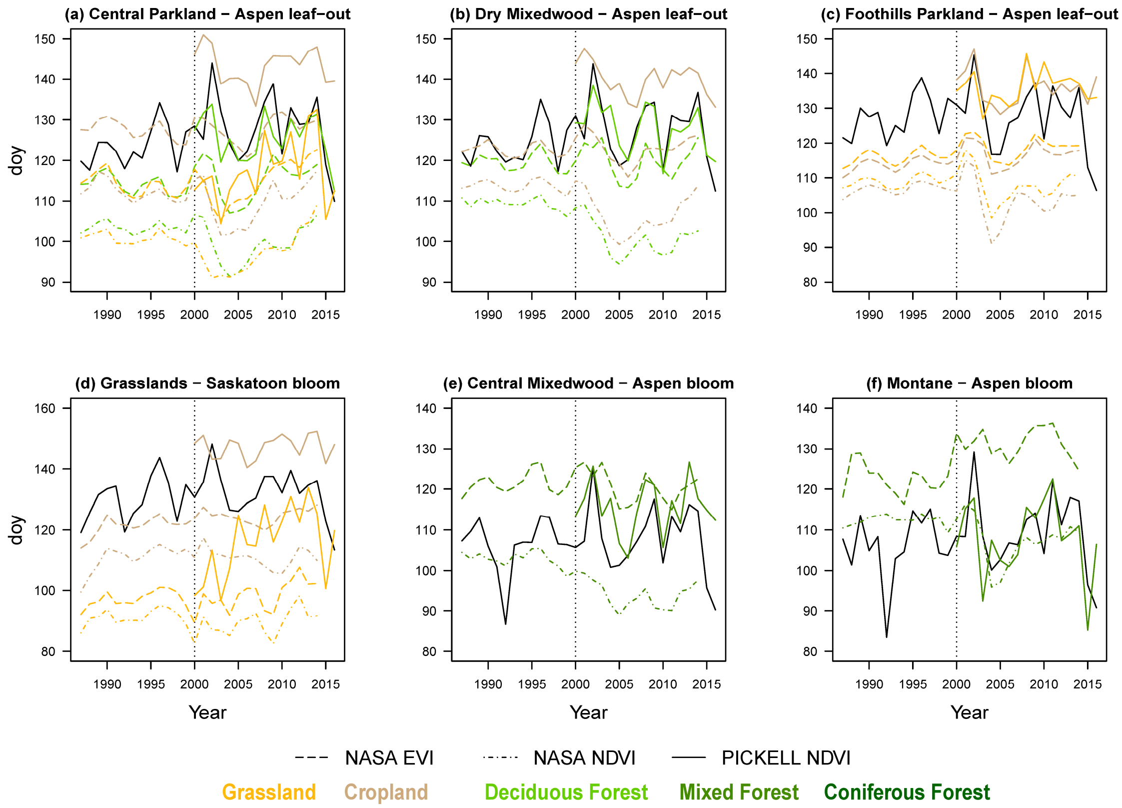

Figure 4. Time series showing the remote sensing green-up dates for the three different remote sensing datasets in each ecoregion by landcover type (colour coded). The black line is the estimated date for a selected phenophase in each ecoregion (as noted in the title of each graph). The vertical dashed line is the year when the NASA data switched from AVHRR to MODIS sensors (2000).

When visualizing the accuracy of remote sensing products, it becomes apparent that all three datasets lack the interannual variability of ground phenology (Figure 5). The slope of the linear regression is always less than 1, and ranges from 0.22 to 0.65 for the relationships presented in Figure 5. The slopes for Pickell NDVI tend to be greater than those for either NASA dataset, affirming that this dataset better captures the interannual variability of phenology. Some of the relationships displayed for the NASA AVHRR and MODIS data have very shallow slopes, but very small residuals around the linear regression line (e.g. Figure 5c; d).

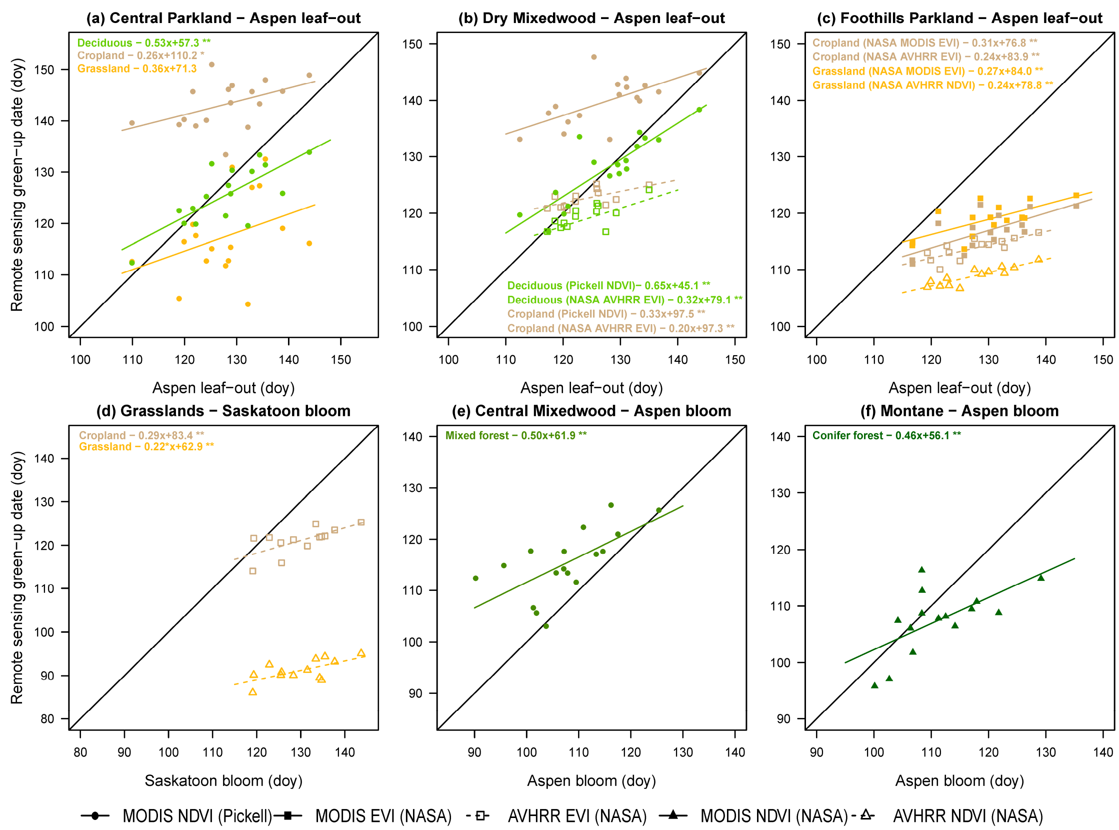

Figure 5. Scatter plots of the remote sensing green-up date in relation to the estimated ground phenology date for the strongest correlated land surface phenology data for each landcover type by ecoregion (black line is the 1:1 line, coloured lines are the linear regression, the linear regression equation and significance (* p<0.05; ** p<0.01) are also included).

3.4 Green-up timing

The remote sensing dataset and the landcover type have considerable effects on the timing of the green-up estimate (Figure 4; Table 3). Average green-up dates range from over two weeks prior to aspen bloom to roughly three weeks after aspen leaf-out (Table 3). The Pickell NDVI dataset has quite variable lags with green-up being as early as aspen bloom in the Montane region, to over 3 weeks after aspen leaf-out for cropland landcover in the Grasslands region. The NASA NDVI dataset has high negative lags relative to aspen leaf-out, and moderately negative to slightly positive lags relative to aspen bloom. In almost all ecoregion – landcover combinations, the NASA EVI dataset has negative lags relative to aspen leaf-out, and positive lags relative to aspen bloom.

Table 3. Average lag in days between the remotely sensed green-up date minus the observed date for aspen first bloom and leaf-out. A negative lag indicates that the green-up estimate precedes the ground observed date for that phenophase.

	MODIS EVI (NASA)		MODIS NDVI (NASA)		MODIS NDVI (Pickell)		AVHRR EVI (NASA)		AVHRR NDVI (NASA)	
Ecoregion/landcover	**Bloom**	**Leaf-out**	**Bloom**	**Leaf-out**	**Bloom**	**Leaf-out**	**Bloom**	**Leaf-out**	**Bloom**	**Leaf-out**
Central Parkland										
Deciduous	6.6	-14.2	-9.3	-30.2	18.5	-2.4	12.2	-9.4	1.5	-20.1
Cropland	18.9	-1.9	0.9	-19.9	36.2	15.4	25.7	4.2	10.4	-11.2
Grassland	6.1	-14.7	-11.5	-32.3	10.4	-10.4	12.1	-9.5	-1.0	-22.6
Dry Mixedwood										
Deciduous	9.1	-9.2	-10.0	-28.3	18.8	0.3	15.1	-4.4	5.0	-14.5
Cropland	12.3	-5.9	-4.0	-22.3	30.8	12.3	18.2	-1.3	9.3	-10.2
Foothills Parkland										
Cropland	15.7	-12.8	2.3	-26.2	37.2	8.7	14.8	-13.6	7.8	-20.6
Grassland	17.6	-10.8	6.9	-21.6	37.1	8.6	16.9	-11.5	9.8	-18.5
Grasslands										
Cropland	17.6	-2.6	4.0	-16.3	42.5	22.3	18.3	-2.0	8.1	-12.3
Grassland	-8.6	-28.9	-17.7	-37.9	11.2	-9.0	-5.5	-25.9	-11.6	-31.9
Central Mixedwood										
Mixed forest	11.2	-11.2	-15.4	-37.7	7.8	-14.5	15.0	-7.0	-3.5	-25.4
Montane										
Conifer forest	20.0	-3.1	-3.8	-26.9	-1.3	-24.3	16.5	-6.5	6.5	-16.5

NASA’s NDVI based green-up estimates precede the EVI estimates in all ecoregions and landcover types (Figure 4). The Pickell NDVI estimates tend to be the latest, expect for the Central Mixedwood and Montane regions (Figure 4e; 4f). Satellite observed green-up is considerably later on cropland than grasslands or deciduous forests, except for the Foothills Parkland region where cropland and grasslands have similar green-up dates. This effect is also highlighted in Figure 2, as the later green-up for cropland in the southern half of the province can be distinguished from the earlier green-up on grasslands to the south-east and forests to the north and west.

Relative to the chosen phenophases, green-up estimates for forested ecosystems are largely unbiased as they centre near the 1:1 line (Figure 5a; b; e; f). Green-up on grassland landcover is estimated early by remote sensing products relative to these phenophases, as the estimates fall well below the 1:1 line (Figure 5a; c; d). Bias for cropland is far more variable, with green-up occurring earlier than the chosen phenophase in some regions (Figure 5a; b), and later than the chosen phenophase in others (Figure 5c; d).

The change from AVHRR to MODIS sensors for the two NASA datasets causes a systematic shift in green-up in several ecoregions that does not correspond to a shift in ground phenology, confirming the need to assess these data separately. The NASA NDVI green-up dates shift earlier in several regions after 2000, which is most evident in the Dry and Central Mixedwood regions (Figure 4b, 4e). This results in the average lags changing by 10 to 14 days (Table 3). Green-up shifts are less evident for the NASA EVI dataset, and the average lags only change by as much as 7 days (Table 3).

4.0 DISCUSSION

4.1 Land surface phenology: performance and implications for Alberta landscapes

The three land surface phenology products have varying correlations with ground phenology, and the timing of green-up ranges drastically between datasets and landcover types. Vegetation indices change in response to anything that changes surface reflectance (Helman 2018; Liang et al. 2011), meaning that green-up can be related to the phenology of many species. Though flowering of most PlantWatch species is not expected to be captured by satellites, the cumulative leafing and blooming of many species is related to temperature accumulation (Beaubien and Hamann 2011b; Menzel et al. 2006); thus, it's unsurprising that green-up is correlated with a wide range of phenophases, including those that are not expected to influence surface reflectance (Appendix Table A2) (Delbart et al. 2015; Misra et al. 2016). Of the phenophases included in this study, aspen bloom and leaf-out are likely to have the greatest influence on satellite observed green-up. Female aspen trees grow long green catkins for several weeks after pollination (Delbart et al. 2015), meaning they could appear green to the satellite prior to leafing. The leafing of shrubs and other tree species would also influence satellite signals. However, there are insufficient leaf-out records in the PlantWatch data for species other than aspen to include these in this analysis.

Assessing the timing of green-up relative to aspen bloom and leaf-out is useful, as aspen bloom is one of the earliest occurring phenophases in Alberta (Figure 3) (Beaubien and Hamann 2011b), and aspen leaf-out is a useful indicator of the onset of photosynthesis for deciduous trees. Land surface phenology estimates of primary productivity are calculated as a proxy of the summed NDVI through the growing season (Park et al. 2016), meaning that biased estimates of spring start would bias productivity estimates. Green-up estimates with high RMSE (i.e. significant bias) can be corrected if the correlation with relevant ground phenophases is strong, while weak correlations indicate that the green-up estimates do not reflect ground phenology in that ecosystem.

4.1.1 Deciduous forests

The strongest correlations between remote sensing green-up dates and ground phenology tend to occur on deciduous forest landscapes. In particular, the Pickell NDVI green-up dates show very strong agreement with aspen leaf-out for deciduous forests, with almost no bias (Table 2; Table 3; Figure 5). These are stronger correlations and lower RMSE than the results of other studies (Delbart et al. 2015; Delbart et al. 2005; White et al. 2009), which may be due to aggregating by ecoregions rather than comparing ground observations with green-up dates of corresponding pixels. The half-maximum SOS retrieval method used by the Pickell NDVI data typically corresponds to the initial leafing of the canopy (Misra et al. 2016; White et al. 1997), as is demonstrated here. Given that aspen is the dominant deciduous tree species in much of Alberta (Natural Regions Committee 2006), it's unsurprising that its leafing would strongly influence satellite observed green-up. The strong correlation and lack of bias between this dataset's green-up dates and aspen leaf-out indicates that the half-maximum SOS retrieval method for NDVI is suitable to predict the onset of canopy photosynthesis for deciduous forests in Alberta.

The NASA EVI and NDVI data have weaker correlations for deciduous forests, and the green-up dates for these datasets occur prior to aspen leaf-out (Table 2, Table 3). The NASA NDVI dataset has green-up dates prior to aspen bloom, suggesting that snowmelt is too strong an influence on green-up detection of this dataset. The weak correlations and high RMSE indicate that the NASA NDVI dataset is unsuitable to predict spring start for Alberta deciduous forests. The Enhanced Vegetation Index is less sensitive than NDVI to light snow cover and bare soil (Helman 2018), which explains why it has later green-up than the NASA NDVI dataset. The NASA MODIS EVI green-up dates for deciduous forests have reasonable correlations with aspen leaf-out, though with an early bias of 9-14 days (Table 2, Table 3). This suggests that this dataset's green-up could be influenced by the progressive growth of catkins on female aspen trees following pollination, or by initial bud expansion and bud burst of aspen leaves, which would occur prior to the recorded leaf-out date. Though this dataset estimates green-up 9-14 days earlier than aspen leaf-out, the strength of the correlation suggests that this bias could be corrected using linear regression to predict the onset of canopy photosynthesis with reasonable accuracy.

4.1.2 Mixed and coniferous forests

Correlations are more variable among the datasets for mixed and coniferous forests, where detecting spring start is more difficult due to the lower annual variability in vegetation index values (Delbart et al. 2005; Hmimina et al. 2013; Jönsson et al. 2010; Karkauskaite et al. 2017). For mixed forests in the Central Mixedwood region, the Pickell NDVI dataset has reasonably strong correlations with aspen bloom (Table 3), and the NASA EVI dataset has similarly strong correlations with aspen leaf-out for both the AVHRR (r: 0.69) and MODIS (r: 0.56) estimates (Appendix Table A2). Similar to deciduous forests, the NASA EVI dataset detects green-up 11 days prior to aspen leaf-out on average (Table 3). However, the Pickell NDVI dataset detects green-up 14 days prior to aspen leaf-out on average in mixed forests, whereas green-up was coincident with leaf-out for deciduous forests (Table 3). The timing of green-up for both datasets suggests that the estimate is related to the growth of green catkins following pollination of female aspen trees, or initial bud expansion and budburst of aspen trees prior to full leaf-out. While deciduous trees don’t begin photosynthesising until they have leafed-out, coniferous trees in mixed forests begin photosynthesising as soon as air temperatures permit, resulting in a longer photosynthetically active period (D’Odorico et al. 2015). Though unsuitable to detect photosynthetic activity of coniferous trees, the strength of these correlations suggests that the NASA EVI and Pickell NDVI datasets are suitable to predict ground phenology and the beginning of photosynthesis for deciduous trees in boreal mixed forests. Linear regression between green-up dates and aspen leaf-out could be used to solve the early bias, to better reflect the onset of photosynthesis for deciduous trees.

Both MODIS NDVI datasets have moderately strong correlations with aspen bloom with little bias for coniferous forests in the Montane region (Table 2, Table 3). The similarity in green-up timing for these datasets is surprising, given that the Pickell NDVI green-up dates tend to be three or more weeks later than the NASA NDVI estimates in other ecoregions. The NASA EVI green-up dates are not correlated well with any phenophase in this region (Appendix Table A2), confirming that EVI is unsuitable for coniferous forests (Liu et al. 2016; Shen et al. 2014; Wu et al. 2014). The results for the MODIS NDVI datasets appear to be an improvement on other studies that found weak or no relationship between NDVI green-up estimates and ground phenology or photosynthesis data for coniferous forests (Jönsson et al. 2010; Liu et al. 2016; Wu et al. 2014). Evergreen trees begin photosynthesising in spring without any change in greenness (D'Odorico et al. 2015; Jönsson et al. 2010; Reed et al. 1994), which makes NDVI and EVI unsuitable to detect photosynthesis in coniferous forests. Other vegetation indices such as the Plant Phenology Index (PPI) (Jin and Eklundh 2014) and Chlorophyll/Carotenoid Index (CCI) (Gamon et al. 2016) are superior for detecting photosynthesis in coniferous forests. While aspen bloom is not expected to influence satellite observed green-up of coniferous forests, the strength of these correlations and lack of bias suggests that NDVI is suitable for predicting the earliest phenophases in this landcover type. This may be influenced by the greening of the understorey, though may also be due to snowmelt.

4.1.3 Cropland

Green-up for cropland is generally later than surrounding forests and grasslands, which other North American studies also found (Zhang et al. 2006; Zhang et al. 2017). There are surprisingly strong correlations between cropland green-up and ground phenology in several ecoregions (Table 2), despite satellite observed green-up of croplands being more related to crop type and management than natural vegetation (Zhang et al. 2006). The Pickell NDVI and NASA EVI datasets have fairly strong correlations with aspen leaf-out for cropland in the Dry Mixedwood region, and both NASA datasets have strong correlations with aspen leaf-out for cropland in the Foothills Parkland region. The NASA AVHRR data also have fairly strong correlations with Saskatoon bloom for cropland in the Grasslands region. The strength of these correlations may be due to aggregating by ecoregions, as another study found little to no relationship between PlantWatch data from cropland landscapes and green-up dates of the associated pixels (Delbart et al. 2015). Given that croplands cover a substantial portion of the earth's surface, monitoring phenology in these regions is crucial to accurately assess changes in phenology. Individual pixels in a cropland matrix have high interannual variability and noise, as different crop types have different green-up timing (Zhang et al. 2017). These results suggest that taking regional averages of cropland green-up can reduce this noise so that green-up predicts regional averages of ground phenology with reasonable accuracy. This could be a useful strategy for researchers wishing to include regions dominated by cropland in land surface phenology studies.

4.1.4 Grasslands

Correlations are most variable among datasets for grassland landcover, which may be due to the PlantWatch species responding differently to climatic drivers than the vegetation that drives satellite observed green-up of Alberta grasslands. Start of growth in the Grasslands region is heavily dependent on preceding water balance, with precipitation events advancing growth, and drought delaying growth (Cui et al. 2017). Li and Guo (2012) demonstrated that NDVI is more strongly correlated with accumulated precipitation than accumulated temperature in Canadian prairie grasslands. Conversely, PlantWatch species' first bloom dates correlate very strongly with accumulated temperature models, with no consideration of precipitation (Beaubien and Hamann 2011b).

To be suitable for citizen science, species included in the PlantWatch dataset must be easily identifiable and have distinct and relatively short phenophases (Beaubien and Hamann 2011a). The grass species that would drive satellite observed green-up in Alberta grasslands lack these criteria, and thus no grass species were chosen for PlantWatch. Shrubland and forests where satellite observed green-up would be more reflective of PlantWatch data, are restricted to moist areas and river valleys in the Grasslands region (Natural Regions Committee 2006). No dataset has strong correlations with ground phenology for grasslands in the Central Parkland region, which may be due to similar factors. The Foothills Parkland has more spread-out aspen forest and shrubland in low-lying areas and north facing slopes (Natural Regions Committee 2006), which may explain why there are stronger correlations with ground phenology for grassland landcover in that ecoregion. It's unclear why there are considerably better correlations with AVHRR than MODIS data for grassland landcover in the Grasslands region (Table 2), though the RMSE of over 30 days brings the validity of these findings into question. The Alberta PlantWatch data are not appropriate to evaluate the suitability of these datasets to track phenology in the grasslands of Alberta, and other sources of ground phenology data should be used to test this.

4.2 Importance of using region specific land surface phenology methods

Helman (2018) recommends that the greenness proxy and method used to extract the phenological metrics should be biome specific, based on knowledge of the dominant plant species, and should be guided by ground observations and comparison of different methods. Land surface phenology studies are routinely carried out at continental or global scales using a single combination of vegetation index, data smoothing technique, and metric retrieval method (Garonna et al. 2015; Jeong et al. 2011; Park et al. 2016; Running et al. 2004; Stöckli and Vidale 2004). However, there is no consensus on a superior method for deriving land surface phenology metrics in global applications (Atkinson et al. 2012; Buitenwerf et al. 2015), and alternative methods perform differently in different vegetation types (Buitenwerf et al. 2015; White et al. 2009). Using a global dataset without considering its suitability for a study region could create false conclusions regarding the timing of phenological metrics and associated measures of productivity. Conversely, applying region specific thresholds to a global study may result in misinterpretations for regions that are not well reflected by the chosen method. Comparison against ground phenology data is necessary to validate the suitability of land surface phenology methods for different regions and vegetation types, in order to avoid these issues.

The two NASA global datasets used in this study do not correlate well with ground phenology in several regions, suggesting that they are not ideal for a northern study region. This is particularly evident for the NASA NDVI dataset, which had weak correlations and green-up dates earlier than aspen bloom in most ecoregions. The beginning of NDVI increase in northern forests is due to snowmelt exposing bare soil and leaf litter, which causes the index to increase (Moulin et al. 1997; Reed et al. 1994; White et al. 2005). It's likely that snowmelt was too strong an influence on green-up detection of this dataset, demonstrating that a 35% retrieval threshold for NDVI is low for most northern ecosystems. Though the NASA EVI dataset has stronger correlations with ground phenology, green-up dates are prior to aspen leaf-out. This suggests that a higher threshold for EVI SOS retrieval would be superior to accurately predict the beginning of photosynthesis in northern ecosystems.

The half-maximum method (50% threshold) used by the Pickell NDVI dataset better predicts the onset of photosynthesis in deciduous forest ecosystems. The original application of the Pickell NDVI dataset was to evaluate the potential of land surface phenology to provide an early warning system for spring fire risk in the forested region of Alberta (Pickell et al. 2017). Though this dataset was not developed with an emphasis on a specific phenophase, it has a strong and unbiased relationship with aspen leaf-out for deciduous forests in Alberta. This dataset also has strong correlations with ground phenology in mixed and coniferous forests, verifying that this green-up threshold accurately predicts spring start in the ecosystems the dataset was designed for. This demonstrates that data processing and metric retrieval methods specific to the region or vegetation type of interest can result in more accurate phenology predictions than those derived from a pre-processed global dataset. Lowering the threshold to derive phenological metrics, such as the 35% threshold used by the NASA datasets, results in earlier SOS dates and later end of season dates (White et al. 1997). Using these lower thresholds for green-up in northern landscapes would thus overestimate growing season length and primary productivity. Researchers should exercise caution when using global datasets and should test against ground data whenever possible in order to avoid mismatches between land surface phenology metrics and relevant phenophases.

4.3 Interannual variability: implications for climate monitoring

The land surface phenology datasets in this study consistently underestimate the interannual variability of phenology, which has also been found in other studies (Fisher and Mustard 2007; Peng et al. 2018; White et al. 2009). Interannual variability of green-up seems to depend primarily on the dataset used, and the landcover type. Linear regression slopes are steeper for forested landcover than grassland or cropland, independent of the correlation strength (Figure 5). For example, the relationships displayed for the Foothills Parkland region all have strong correlations with aspen leaf-out ($r > 0.69$), although with quite shallow slopes (<0.31) (Figure 5c). The linear regression slopes are also steeper for the Pickell NDVI dataset than either NASA dataset, and the slopes for the NASA datasets are steeper for MODIS data than AVHRR (Figure 5). Failing to reproduce the interannual variability of phenology from remote sensing risks underestimating the impact of climate change on ecosystems, as the rate of change for spring start, fall senescence, and growing season length would all be underestimated. By comparing with ground phenology data, linear regression equations could be used to convert green-up dates into predicted dates for a ground based phenophase. The equations in Figure 5 for the strongly correlated combinations are a suitable conversion of green-up to predicted dates for the chosen phenophases, which could be used for these datasets in these landscape types. Doing so would increase the interannual variability of remotely sensed time-series, in order to better estimate trends in phenology. Several other factors that influence the variability from remote sensing could also be considered in order to better estimate the variability of phenology.

Landscape heterogeneity, pixel size, and sensor type contribute to the differences in interannual variability between datasets and ecoregions. The extraction points were placed in locations with several clustered pixels of the same landcover type, which results in stronger agreement with ground phenology compared to non-buffered pixels (Doktor et al. 2009). Given that the landcover data are at the same spatial resolution as the Pickell NDVI data (250m), it's fair to expect that this dataset's green-up are from relatively contiguous pixels of the target landcover type. However, the pixel size of the NASA data is roughly 500 times the size of the landcover pixels, and thus it's likely that these pixels include unintended landcover types. Large pixels fail to reveal the fine scale variabilities in climate and vegetation that influence land surface phenology (Fisher and Mustard 2007), which is evident when comparing the Pickell NDVI dataset to the NASA datasets in Figure 2. Additionally, vegetation indices experience a less distinct increase to full summer greenness for heterogeneous pixels (Doktor et al. 2009), which explains why the NASA datasets are less variable than the Pickell NDVI dataset. The NASA data collected by AVHRR sensors have less interannual variability than from MODIS (Figure 4; Figure 5), which may be explained by the better ability of MODIS to diminish the effects of cloud contamination, atmospheric variability, and sensor view angle (Zhang et al. 2003; Zhang et al. 2001). MODIS data are also available at finer resolutions (250m) than AVHRR data (1km) (Helman 2018), furthering this sensors capacity to better capture interannual variability. Using finer resolution imagery and data from newer sensors such as MODIS are appropriate strategies to better estimate the interannual variability of phenology, and better delineate spatial heterogeneity in mixed landscapes.

While reduced interannual variability would underestimate the rate of change for spring start, land surface phenology may also not be sensitive to phenophases that are shifting the most rapidly in northern ecosystems. Significant trends of spring advance in Alberta have only been found for the two earliest phenophases in the Alberta PlantWatch database: aspen and prairie crocus first bloom (Beaubien and Hamann 2011b; Beaubien and Freeland 2000). Thus, green-up estimates that accurately capture these earliest phenophases would be preferable to monitor the impact of climate change on spring start. While green-up dates have acceptable correlations with these phenophases in several ecoregions, correlations tend to be stronger for aspen leaf-out (Appendix Table A2). Snowmelt plays too strong an influence on green-up estimates from lower NDVI thresholds in Alberta, and thus this vegetation index is unsuitable to detect these phenophases. Though EVI is less sensitive to snowmelt, the NASA EVI dataset tends to have weaker correlations with aspen and prairie crocus first bloom than the NDVI datasets (Appendix Table A2). Thus, alternative vegetation indices with reduced sensitivity to snowmelt such as the Normalised Difference Water Index (NDWI) (Delbart et al. 2005), and the Plant Phenology Index (PPI) (Jin and Eklundh 2014) should be evaluated for their potential to detect the early phenophases that are most sensitive to climate change.

5. CONCLUSION

This study confirms previous findings that different remote sensing techniques have variable agreement with ground phenology data. In order to accurately predict the onset of photosynthesis, linear regression can be used for datasets that consistently estimate an early or late green-up relative to canopy leaf-out. This could be used to correct the early bias of green-up estimates for the NASA EVI dataset in deciduous and mixed forests of Alberta, and for the Pickell dataset in mixed forests. The lags in Table 3 and linear regression equations in Figure 5 could form the basis to correct these biases. The strong correlations between green-up and ground phenology for cropland landcover suggest that aggregating data by ecoregions is a suitable method to reduce the noise and heterogeneity in green-up that cause mismatches on this landcover type. Comparing against ground phenology data is essential to ensure land surface phenology accurately predicts the timing of relevant phenophases for a region and vegetation type of interest. Maintaining and expanding ground phenology networks such as PlantWatch is crucial to ensure the continuity of this secondary data source that can complement and validate land surface phenology.

Different vegetation types and ecosystems demonstrate different characteristics of green-up; thus, developing global land surface phenology methods and products is problematic. Using global datasets without considering the suitability for any study region could result in false conclusions regarding the timing and trends of green-up and senescence, and could bias estimates of primary productivity. The two NASA datasets in this study estimate an early green-up relative to canopy leaf-out in Alberta forests, and the NDVI dataset seems more related to snowmelt that green-up in most ecoregions. This suggests that global thresholds for green-up may underestimate spring start in northern ecosystems, which would lead to overestimations of growing season length and primary productivity. The stronger correlations with ground phenology for the Pickell NDVI dataset in forested regions demonstrates that developing a dataset specific to a vegetation type of interest can result in better green-up estimations than those from pre-processed global datasets. Comparisons against ground phenology data should be used when possible to set region specific land surface phenology methods, and to ensure data accurately reflect the vegetation type of any study region.

Land surface phenology universally underestimates the interannual variability of phenology, which risks underestimating the impact of climate change on ecosystems. Interannual variation is better represented for forested landscapes, and fine resolution data substantially increase the variability compared to coarser resolutions. By comparing green-up estimates with ground phenology data, linear regression equations can be developed to convert green-up estimates into predicted dates for a phenophase. This would increase the variability of remotely sensed time-series, in order to better estimate trends in phenology. There should also be a focus on developing remote sensing techniques that accurately detect the earliest phenophases in northern ecosystems, given that these are shifting the most rapidly.

6. LITTERATURE CITED

Ahas, R., Aasa, A., Menzel, A., Fedotova, V.G., and Scheifinger, H. 2002. Changes in European spring phenology. International Journal of Climatology **22**(14): 1727-1738.

Atkinson, P.M., Jeganathan, C., Dash, J., and Atzberger, C. 2012. Inter-comparison of four models for smoothing satellite sensor time-series data to estimate vegetation phenology. Remote Sensing of Environment **123**: 400-417.

Badeck, F.-W., Bondeau, A., Böttcher, K., Doktor, D., Lucht, W., Schaber, J., and Sitch, S. 2004. Responses of spring phenology to climate change. The New Phytologist **162**(2): 295-309.

Beaubien, E., and Hamann, A. 2011a. Plant phenology networks of citizen scientists: recommendations from two decades of experience in Canada. International Journal of Biometeorology **55**(6): 833-841.

Beaubien, E., and Hamann, A. 2011b. Spring flowering response to climate change between 1936 and 2006 in Alberta, Canada. BioScience **61**(7): 514-524.

Beaubien, E.G., and Freeland, H.J. 2000. Spring phenology trends in Alberta, Canada: links to ocean temperature. International Journal of Biometeorology **44**(2): 53-59.

Betancourt, J.L., Schwartz, M.D., Breshards, D.D., Brewer, C.A., Frazer, G., Gross, J.E., Mazer, S.J., Reed, B.C., and Wilson, B.E. 2007. Evolving plans for the USA National Phenology Network. EOS **88**(19): 211-211.

Buitenwerf, R., Rose, L., and Higgins, S.I. 2015. Three decades of multi-dimensional change in global leaf phenology. Nature Climate Change **5**: 364-368.

Butler, D. 2009. asreml: asreml() fits the linear mixed model. R package version 3.0. www.vsni.co.uk

CEC. 2013. 2010 North American Land Cover at 250 m spatial resolution. *Edited by* Natural Resources Canada/ The Canada Centre for Mapping and Earth Observation (NRCan/CCMEO), United States Geological Survey (USGS), and Insituto Nacional de Estadística y Geografía (INEGI) Comisión Nacional para el Conocimiento y Uso de la Biodiversidad (CONABIO) and Comisión Nacional Forestal (CONAFOR). Commission for Environmental Cooperation, Montréal, Québec, Canada.

Cleland, E.E., Chuine, I., Menzel, A., Mooney, H.A., and Schwartz, M.D. 2007. Shifting plant phenology in response to global change. Trends in Ecology & Evolution **22**(7): 357-365.

Collinson, N., and Sparks, T. 2008. Phenology – Nature's calendar: An overview of results from the UK Phenology Network. Arboricultural Journal **30**(4): 271-278.

Cui, T., Martz, L., and Guo, X. 2017. Grassland phenology response to drought in the Canadian prairies. Remote Sensing **9**(12): 1258.

D'Odorico, P., Gonsamo, A., Gough, C.M., Bohrer, G., Morison, J., Wilkinson, M., Hanson, P.J., Gianelle, D., Fuentes, J.D., and Buchmann, N. 2015. The match and mismatch between photosynthesis and land surface phenology of deciduous forests. Agricultural and Forest Meteorology **214-215**: 25-38.

Delbart, N., Beaubien, E., Kergoat, L., and Le Toan, T. 2015. Comparing land surface phenology with leafing and flowering observations from the PlantWatch citizen network. Remote Sensing of Environment **160**: 273-280.

Delbart, N., Kergoat, L., Le Toan, T., Lhermitte, J., and Picard, G. 2005. Determination of phenological dates in boreal regions using normalized difference water index. Remote Sensing of Environment **97**(1): 26-38.

Didan, K., and Barreto, A. 2016a. NASA MEaSUREs Vegetation Index and Phenology (VIP) Phenology EVI2 Yearly Global 0.05Deg CMG. NASA EOSDIS Land Processes DAAC.

Didan, K., and Barreto, A. 2016b. NASA MEaSUREs Vegetation Index and Phenology (VIP) Phenology NDVI Yearly Global 0.05Deg CMG. NASA EOSDIS Land Processes DAAC.

Didan, K., Munoz, A.B., Miura, T., Tsend-Ayush, J., Zhang, X., Friedl, M., Gray, J., van Leeuwen, W., Czapla-Myers, J., Jenkerson, C., Maiersperger, T., and Meyer, D. 2018. Multi-sensor vegetation index and phenology earth science data records: algorithm theoretical basis document and user guide. Available from https://vip.arizona.edu/VIP_ATBD_UsersGuide.php [accessed April 26 2018].

Doktor, D., Bondeau, A., Koslowski, D., and Badeck, F.-W. 2009. Influence of heterogeneous landscapes on computed green-up dates based on daily AVHRR NDVI observations. Remote Sensing of Environment **113**(12): 2618-2632.

Donnelly, A., and Yu, R. 2017. The rise of phenology with climate change: an evaluation of IJB publications. International Journal of Biometeorology **61**(S1): 29-50.

Feldman, R.E., Žemaitė, I., and Miller-Rushing, A.J. 2018. How training citizen scientists affects the accuracy and precision of phenological data. International Journal of Biometeorology **62**(8): 1421-1435.

Fisher, J.I., and Mustard, J.F. 2007. Cross-scalar satellite phenology from ground, Landsat, and MODIS data. Remote Sensing of Environment **109**(3): 261-273.

Fitter, A.H., and Fitter, R.S.R. 2002. Rapid changes in flowering time in British plants. Science **296**(5573): 1689.

Gamon, J.A., Huemmrich, K.F., Wong, C.Y.S., Ensminger, I., Garrity, S., Hollinger, D.Y., Noormets, A., and Peñuelas, J. 2016. A remotely sensed pigment index reveals photosynthetic phenology in evergreen conifers. Proceedings of the National Academy of Sciences **113**(46): 13087-13092.

Garonna, I., Jong, R., and Schaepman Michael, E. 2015. Variability and evolution of global land surface phenology over the past three decades (1982–2012). Global Change Biology **22**(4): 1456-1468.

Goward, S.N., Tucker, C.J., and Dye, D.G. 1985. North American vegetation patterns observed with the NOAA-7 advanced very high resolution radiometer. Vegetatio **64**(1): 3-14.

Gu, L., Post, W.M., Baldocchi, D.D., Black, T.A., Verma, S.B., Vesala, T., and Wofsy, S.C. 2003. Phenology of Vegetation Photosynthesis. *In* Phenology: An Integrative Environmental Science. *Edited by* Mark D. Schwartz. Kluwer Academic Publishers, Dodrecht, Netherlands. pp. 467-485.

Hanes, J.M., Liang, L., and Morisette, J.T. 2014. Land Surface Phenology. *In* Biophysical Applications of Satellite Remote Sensing. *Edited by* Jonathan M. Hanes. Springer, Heidelberg; New York; Dordrecht; London. pp. 99-125.

Helman, D. 2018. Land surface phenology: What do we really ‘see’ from space? Science of the Total Environment **618**: 665-673.

Hmimina, G., Dufrêne, E., Pontailler, J.-Y., Delpierre, N., Aubinet, M., Caquet, B., de Grandcourt, A., Burban, B., Flechard, C., Granier, A., Gross, P., Heinesch, B., Longdoz, B., Moureaux, C., Ourcival, J.-M., Rambal, S., Saint André, L., and Soudani, K. 2013. Evaluation of the potential of MODIS satellite data to predict vegetation phenology in different biomes: An investigation using ground-based NDVI measurements. Remote Sensing of Environment **132**: 145-158.

Huete, A., Didan, K., Miura, T., Rodriguez, E.P., Gao, X., and Ferreira, L.G. 2002. Overview of the radiometric and biophysical performance of the MODIS vegetation indices. Remote Sensing of Environment **83**(1): 195-213.

Huete, A., Miura, T., Yoshioka, H., Ratana, P., and Broich, M. 2014. Indices of Vegetation Activity. *In* Biophysical Applications of Satellite Remote Sensing. *Edited by* Jonathan M. Hanes. Springer, Heidelberg; New York; Dordrecht; London.

Intergovernmental Panel on Climate Change. 2007. Climate Change 2007: Synthesis Report. Contribution of Working Groups I, II, and III to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. Core Writing Team, Pachauri RK, Reisinger A (eds). IPCC, Geneva, Switzerland, 104 pp.

Jeong, S.-J., Ho, C.-H., Gim, H.-J., and Brown Molly, E. 2011. Phenology shifts at start vs. end of growing season in temperate vegetation over the Northern Hemisphere for the period 1982–2008. Global Change Biology **17**(7): 2385-2399.

Jin, H., and Eklundh, L. 2014. A physically based vegetation index for improved monitoring of plant phenology. Remote Sensing of Environment **152**: 512-525.

Jönsson, A.M., Eklundh, L., Hellström, M., Bärring, L., and Jönsson, P. 2010. Annual changes in MODIS vegetation indices of Swedish coniferous forests in relation to snow dynamics and tree phenology. Remote Sensing of Environment **114**(11): 2719-2730.

Karkauskaite, P., Tagesson, T., and Fensholt, R. 2017. Evaluation of the Plant Phenology Index (PPI), NDVI and EVI for start-of-season trend analysis of the northern hemisphere boreal zone. Remote Sensing **9**(5): 485.

Kross, A., Fernandes, R., Seaquist, J., and Beaubien, E. 2011. The effect of the temporal resolution of NDVI data on season onset dates and trends across Canadian broadleaf forests. Remote Sensing of Environment **115**(6): 1564-1575.

Leinonen, I., and Kramer, K. 2002. Applications of phenological models to predict the future carbon sequestration potential of boreal forests. Climatic Change **55**(1): 99-113.

Li, Z., and Guo, X. 2012. Detecting climate effects on vegetation in northern mixed prairie using NOAA AVHRR 1-km time-series NDVI data. Remote Sensing **4**(1): 120-134.

Liang, L., Schwartz, M.D., and Fei, S. 2011. Validating satellite phenology through intensive ground observation and landscape scaling in a mixed seasonal forest. Remote Sensing of Environment **115**(1): 143-157.

Liu, Y., Wu, C., Peng, D., Xu, S., Gonsamo, A., Jassal, R.S., Altaf Arain, M., Lu, L., Fang, B., and Chen, J.M. 2016. Improved modeling of land surface phenology using MODIS land surface reflectance and temperature at evergreen needleleaf forests of central North America. Remote Sensing of Environment **176**: 152-162.

Menzel, A., Sparks, T.H., Estrella, N., Koch, E., Aasa, A., Ahas, R., Alm-Kübler, K., Bissolli, P., Braslavská, O.G., Briede, A., Chmielewski, F.M., Crepinsek, Z., Curnel, Y., Dahl, Å., Defila, C., Donnelly, A., Filella, Y., Jatczak, K., MÅGe, F., Mestre, A., Nordli, Ø., Peñuelas, J., Pirinen, P., Remišová, V., Scheifinger, H., Striz, M., Susnik, A., Van Vliet, A.J.H., Wielgolaski, F.E., Zach, S., and Zust, A.N.A. 2006. European phenological response to climate change matches the warming pattern. Global Change Biology **12**(10): 1969-1976.

Misra, G., Buras, A., and Menzel, A. 2016. Effects of different methods on the comparison between land surface and ground phenology—A methodological case study from south-western Germany. Remote Sensing **8**(9): 753.

Moulin, S., Kergoat, L., Viovy, N., and Dedieu, G. 1997. Global-scale assessment of vegetation phenology using NOAA/AVHRR satellite measurements. Journal of Climate **10**(6): 1154-1170.

Myneni, R.B., Keeling, C.D., Tucker, C.J., Asrar, G., and Nemani, R.R. 1997. Increased plant growth in the northern high latitudes from 1981 to 1991. Nature **386**: 698.

Natural Regions Committee. 2006. Natural Regions and Subregions of Alberta. Government of Alberta. Compiled by D.J. Downing and W.W. Pettapiece. Government of Alberta. Pub. No. T/852.

Park, T., Ganguly, S., Tømmervik, H., Euskirchen, E.S., Høgda, K.-A., Karlsen, S.R., Brovkin, V., Nemani, R.R., and Myneni, R.B. 2016. Changes in growing season duration and productivity of northern vegetation inferred from long-term remote sensing data. Environmental Research Letters **11**(8): 084001.

Peng, D., Wu, C., Zhang, X., Yu, L., Huete, A.R., Wang, F., Luo, S., Liu, X., and Zhang, H. 2018. Scaling up spring phenology derived from remote sensing images. Agricultural and Forest Meteorology **256-257**: 207-219.

Pickell, P.D., Coops, N.C., Ferster, C.J., Bater, C.W., Blouin, K.D., Flannigan, M.D., and Zhang, J. 2017. An early warning system to forecast the close of the spring burning window from satellite-observed greenness. Scientific Reports **7**: 14190.

Pouliot, D., Latifovic, R., Fernandes, R., and Olthof, I. 2011. Evaluation of compositing period and AVHRR and MERIS combination for improvement of spring phenology detection in deciduous forests. Remote Sensing of Environment **115**(1): 158-166.

R Core Team. 2013. A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Reed, B.C., Brown, J.F., Vanderzee, D., Loveland, T.R., Merchant, J.W., and Ohlen, D.O. 1994. Measuring phenological variability from satellite imagery. Journal of Vegetation Science **5**: 703-714.

Reed, B.C., White, M., and Brown, J.F. 2003. Remote Sensing Phenology. *In* Phenology: An Integrative Environmental Science. *Edited by* Mark D. Schwartz. Kluwer Academic Publishers, Dordrecht, Netherlands. pp. 365-381.

Richardson, A.D., Keenan, T.F., Migliavacca, M., Ryu, Y., Sonnentag, O., and Toomey, M. 2013. Climate change, phenology, and phenological control of vegetation feedbacks to the climate system. Agricultural and Forest Meteorology **169**: 156-173.

Root, T.L., Price, J.T., Hall, K.R., Schneider, S.H., Rosenzweig, C., and Pounds, J.A. 2003. Fingerprints of global warming on wild animals and plants. Nature **421**: 57-60.

Rosenzweig, C., Casassa, G., Karoly, D.J., Imeson, A., Liu, C., Menzel, A., Rawlins, S., Root, T.L., Seguin, B., and Tryjanowski, P. 2007. Assessment of observed changes and responses in natural and managed systems. *In* Climate Change 2007: Impacts, Adaptation and Vulnerability. Contribution of Working Group II to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. *Edited by* M. L. Parry, O. F. Canziani, J.P. Palutikof, P. J. van der Linden, and C. E. Hanson. Cambridge University Press, Cambridge, UK. pp. 79-131.

Running, S.W., Nemani, R.R., Heinsch, F.A., Zhao, M., Reeves, M., and Hashimoto, H. 2004. A Continuous Satellite-Derived Measure of Global Terrestrial Primary Production. BioScience **54**(6): 547-560.

Schloss, A.L., Kicklighter, D.W., Kaduk, J., and Wittenberg, U. 2001. Comparing global models of terrestrial net primary productivity (NPP): comparison of NPP to climate and the Normalized Difference Vegetation Index (NDVI). Global Change Biology **5**(S1): 25-34.

Schwartz, M.D., Ahas, R., and Anto, A. 2006. Onset of spring starting earlier across the Northern Hemisphere. Global Change Biology **12**(2): 343-351.

Schwartz, M.D., Reed, B.C., and White, M.A. 2002. Assessing satellite-derived start-of-season measures in the conterminous USA. International Journal of Climatology **22**(14): 1793-1805.

Schwartz, M.D., and Reiter, B.E. 2000. Changes in North American spring. International Journal of Climatology **20**: 929-932.

Shen, M., Tang, Y., Desai, A.R., Gough, C., and Chen, K. 2014. Can EVI-derived land-surface phenology be used as a surrogate for phenology of canopy photosynthesis? International Journal of Remote Sensing **35**(3): 1162-1174.

Sparks, T., Menzel, A., and Chr, N. 2009. European cooperation in plant phenology. Climate Research **39**: 175-177.

Stöckli, R., and Vidale, P.L. 2004. European plant phenology and climate as seen in a 20-year AVHRR land-surface parameter dataset. International Journal of Remote Sensing **25**(17): 3303-3330.

Tucker, C.J. 1979. Red and photographic infrared linear combinations for monitoring vegetation. Remote Sensing of Environment **8**(2): 127-150.

Tucker, C.J., Slayback, D.A., Pinzon, J.E., Los, S.O., Myneni, R.B., and Taylor, M.G. 2001. Higher northern latitude normalized difference vegetation index and growing season trends from 1982 to 1999. International Journal of Biometeorology **45**(4): 184-190.

White, K., Pontius, J., and Schaberg, P. 2014. Remote sensing of spring phenology in northeastern forests: A comparison of methods, field metrics and sources of uncertainty. Remote Sensing of Environment **148**: 97-107.

White, M., A., De Beurs, K.M., Didan, K., Inouye, D.W., Richardson, A.D., Jensen, O.P., O'Keefe, J., Zhang, G., Nemani, R.R., Van Leeuwen, W.J.D., Brown Jesslyn, F., De Wit, A., Schaepman, M., Lin, X., Dettinger, M., Bailey, A.S., Kimball, J., Schwartz, M.D., Baldocchi, D.D., Lee, J.T., and Lauenroth, W.K. 2009. Intercomparison, interpretation, and assessment of spring phenology in North America estimated from remote sensing for 1982–2006. Global Change Biology **15**(10): 2335-2359.

White, M.A., Hoffman, F., Hargrove, W.W., and Nemani, R.R. 2005. A global framework for monitoring phenological responses to climate change. Geophysical Research Letters **32**(4): L04705.

White, M.A., Running, S.W., and Thornton, P.E. 1999. The impact of growing-season length variability on carbon assimilation and evapotranspiration over 88 years in the eastern US deciduous forest. International Journal of Biometeorology **42**(3): 139-145.

White, M.A., Thornton, P.E., and Running, S.W. 1997. A continental phenology model for monitoring vegetation responses to interannual climatic variability. Global Biogeochemical Cycles **11**(2): 217-234.

Willmott, C.J. 1982. Some comments on the evaluation of model performance. Bulletin American Meteorological Society **63**(22): 1309-1313.

Wu, C., Gonsamo, A., Gough, C.M., Chen, J.M., and Xu, S. 2014. Modeling growing season phenology in North American forests using seasonal mean vegetation indices from MODIS. Remote Sensing of Environment **147**: 79-88.

Xiangming, X., Cui, J., and Dong, J. 2014. Gross Primary Production of Terrestrial Vegetation. *In* Biophysical Applications of Satellite Remote Sensing. *Edited by* Jonathan M. Hanes. Springer, Heidelberg; New York; Dordrecht; London. pp. 127-148.

Zhang, X., Friedl, M.A., and Schaaf, C.B. 2006. Global vegetation phenology from Moderate Resolution Imaging Spectroradiometer (MODIS): Evaluation of global patterns and comparison with in situ measurements. Journal of Geophysical Research: Biogeosciences **111**: G04017.

Zhang, X., Friedl, M.A., Schaaf, C.B., Strahler, A.H., Hodges, J.C.F., Gao, F., Reed, B.C., and Huete, A. 2003. Monitoring vegetation phenology using MODIS. Remote Sensing of Environment **84**(3): 471-475.

Zhang, X., Hodges, J.C.F., Schaaf, C.B., Friedl, M.A., Strahler, A.H., and Gao, F. 2001. Global vegetation phenology from AVHRR and MODIS data. *In* IGARSS 2001. Scanning the Present and Resolving the Future. Proceedings. IEEE 2001 International Geoscience and Remote Sensing Symposium (Cat. No.01CH37217). pp. 2262-2264.

Zhang, X., Wang, J., Gao, F., Liu, Y., Schaaf, C., Friedl, M., Yu, Y., Jayavelu, S., Gray, J., Liu, L., Yan, D., and Henebry, G.M. 2017. Exploration of scaling effects on coarse resolution land surface phenology. Remote Sensing of Environment **190**: 318-330.

Zhou, L., Tucker, C.J., Kaufmann, R.K., Slayback, D., Shabanov, N.V., and Myneni, R.B. 2001. Variation in northern vegetation activity inferred from satellite data of vegetation index during 1981 to 1999. Journal of Geophysical Research **106**: 20069-20083.

APPENDIX

Appendix Table A1. Number of Alberta PlantWatch observations by species for the six ecoregions in this study.

	Ecoregion					
Species	Central Parkland	Dry Mixedwood	Foothills Parkland	Grasslands	Central Mixedwood	Montane
Saskatoon	2,140	1,634	447	650	282	286
Early blue violet	1,647	1,237	416	474	201	344
Prairie crocus	1,532	433	692	836	72	382
Aspen poplar	1,824	1,250	247	250	158	137
Chokecherry	1,639	1,056	176	494	152	137
Northern bedstraw	1,291	1,194	315	313	222	253
Golden bean	1,707	209	605	877	4	162
Yarrow	1,159	1,039	309	442	264	276
Star-flowered Solomon's seal	1,353	714	382	446	89	267

Appendix Table A2. Correlation coefficients (r) between the remote sensing green-up dates and the first bloom of all species, as well as leaf-out for aspen. Stronger correlations are highlighted in red. Species are sorted from the earliest to latest (PC: prairie crocus; AP1: aspen bloom; AP2: aspen leaf-out; EV: early blue violet; GB: golden bean; SK: Saskatoon; SS: Star-flowered Solomon's seal; NB: northern bedstraw; Y: yarrow).

				Species									
Ecoregion	Sensor type	Data and VI	Landcover	PC	AP1	AP2	EV	GB	SK	CC	SS	NB	Y
Central Parkland	MODIS	NASA EVI	Deciduous forest	0.49	0.52	0.56	0.54	0.38	0.56	0.46	0.58	0.59	0.74
Central Parkland	MODIS	NASA EVI	Cropland	0.33	0.30	0.42	0.46	0.35	0.54	0.45	0.61	0.77	0.78
Central Parkland	MODIS	NASA EVI	Grassland	0.10	0.18	0.11	0.10	-0.09	0.11	0.09	0.19	0.48	0.32
Central Parkland	MODIS	NASA NDVI	Deciduous forest	0.25	0.31	0.30	0.21	0.04	0.24	0.21	0.29	0.35	0.37
Central Parkland	MODIS	NASA NDVI	Cropland	0.23	0.32	0.32	0.24	0.07	0.29	0.27	0.35	0.51	0.51
Central Parkland	MODIS	NASA NDVI	Grassland	0.15	0.21	0.18	0.16	0.01	0.13	0.14	0.17	0.37	0.16
Central Parkland	MODIS	Pickell NDVI	Deciduous forest	0.68	0.72	0.75	0.73	0.62	0.74	0.71	0.76	0.72	0.82
Central Parkland	MODIS	Pickell NDVI	Cropland	0.47	0.50	0.46	0.47	0.32	0.46	0.39	0.47	0.53	0.69
Central Parkland	MODIS	Pickell NDVI	Grassland	0.42	0.50	0.37	0.38	0.20	0.31	0.25	0.30	0.35	0.37
Central Parkland	AVHRR	NASA EVI	Deciduous forest	0.08	0.12	0.26	0.21	0.30	0.38	0.36	0.26	0.26	0.01
Central Parkland	AVHRR	NASA EVI	Cropland	0.11	-0.03	0.24	0.21	0.26	0.31	0.36	0.29	0.32	0.06
Central Parkland	AVHRR	NASA EVI	Grassland	0.10	0.14	0.22	0.17	0.24	0.34	0.28	0.23	0.16	-0.04
Central Parkland	AVHRR	NASA NDVI	Deciduous forest	0.20	0.11	0.37	0.29	0.38	0.47	0.40	0.38	0.37	0.25
Central Parkland	AVHRR	NASA NDVI	Cropland	0.08	0.09	0.16	0.08	0.14	0.25	0.24	0.19	0.15	-0.05
Central Parkland	AVHRR	NASA NDVI	Grassland	0.40	0.38	0.40	0.40	0.35	0.46	0.38	0.34	0.27	0.18
Dry Mixedwood	MODIS	NASA EVI	Deciduous forest	0.43	0.45	0.64	0.67	0.67	0.62	0.50	0.56	0.58	0.61
Dry Mixedwood	MODIS	NASA EVI	Cropland	0.45	0.44	0.57	0.57	0.53	0.53	0.46	0.53	0.70	0.84
Dry Mixedwood	MODIS	NASA NDVI	Deciduous forest	0.42	0.31	0.52	0.43	0.48	0.42	0.34	0.40	0.62	0.67
Dry Mixedwood	MODIS	NASA NDVI	Cropland	0.43	0.35	0.50	0.45	0.45	0.42	0.36	0.44	0.67	0.80
Dry Mixedwood	MODIS	Pickell NDVI	Deciduous forest	0.71	0.75	0.86	0.80	0.81	0.83	0.79	0.70	0.70	0.62
Dry Mixedwood	MODIS	Pickell NDVI	Cropland	0.52	0.58	0.63	0.60	0.59	0.58	0.53	0.51	0.68	0.75
Dry Mixedwood	AVHRR	NASA EVI	Deciduous forest	0.54	0.61	0.73	0.69	0.67	0.79	0.66	0.57	0.67	0.63
Dry Mixedwood	AVHRR	NASA EVI	Cropland	0.35	0.50	0.64	0.56	0.65	0.77	0.59	0.57	0.57	0.53
Dry Mixedwood	AVHRR	NASA NDVI	Deciduous forest	0.19	0.32	0.29	0.28	0.21	0.31	0.24	0.19	0.31	0.24
Dry Mixedwood	AVHRR	NASA NDVI	Cropland	0.37	0.51	0.56	0.47	0.51	0.68	0.48	0.44	0.50	0.45
Foothills Parkland	MODIS	NASA EVI	Cropland	0.68	0.65	0.71	0.81	0.75	0.76	0.62	0.73	0.62	0.73
Foothills Parkland	MODIS	NASA EVI	Grassland	0.65	0.60	0.69	0.79	0.73	0.78	0.65	0.77	0.69	0.76
Foothills Parkland	MODIS	NASA NDVI	Cropland	0.70	0.66	0.66	0.73	0.66	0.58	0.53	0.58	0.43	0.66
Foothills Parkland	MODIS	NASA NDVI	Grassland	0.69	0.64	0.64	0.71	0.67	0.56	0.49	0.56	0.41	0.62
Foothills Parkland	MODIS	Pickell NDVI	Cropland	0.40	0.40	0.47	0.44	0.44	0.47	0.45	0.44	0.41	0.63
Foothills Parkland	MODIS	Pickell NDVI	Grassland	0.30	0.25	0.33	0.37	0.29	0.43	0.40	0.45	0.51	0.65
Foothills Parkland	AVHRR	NASA EVI	Cropland	0.42	0.46	0.77	0.75	0.78	0.79	0.74	0.76	0.76	0.77
Foothills Parkland	AVHRR	NASA EVI	Grassland	0.47	0.51	0.81	0.79	0.81	0.82	0.78	0.77	0.80	0.74
Foothills Parkland	AVHRR	NASA NDVI	Cropland	0.43	0.45	0.75	0.78	0.77	0.76	0.71	0.70	0.71	0.75
Foothills Parkland	AVHRR	NASA NDVI	Grassland	0.66	0.67	0.89	0.89	0.86	0.87	0.87	0.79	0.77	0.73
Grasslands	MODIS	NASA EVI	Cropland	0.26	0.26	0.11	0.19	0.18	0.14	-0.06	0.23	0.23	0.24
Grasslands	MODIS	NASA EVI	Grassland	0.21	0.20	0.01	-0.03	0.09	0.00	-0.15	0.01	-0.06	0.14
Grasslands	MODIS	NASA NDVI	Cropland	0.22	0.13	0.02	0.02	0.04	-0.02	-0.26	0.01	-0.31	-0.03
Grasslands	MODIS	NASA NDVI	Grassland	0.15	0.10	-0.05	-0.12	0.00	-0.07	-0.20	-0.09	-0.12	0.10
Grasslands	MODIS	Pickell NDVI	Cropland	0.02	0.09	0.06	0.07	0.01	0.09	0.09	0.20	0.43	0.41
Grasslands	MODIS	Pickell NDVI	Grassland	0.21	0.27	0.17	0.09	0.13	0.12	0.14	0.21	0.27	0.40
Grasslands	AVHRR	NASA EVI	Cropland	0.21	0.14	0.58	0.41	0.58	0.68	0.64	0.70	0.64	0.78
Grasslands	AVHRR	NASA EVI	Grassland	0.40	0.35	0.55	0.38	0.53	0.61	0.57	0.54	0.34	0.65
Grasslands	AVHRR	NASA NDVI	Cropland	0.20	0.07	0.55	0.40	0.55	0.66	0.61	0.68	0.58	0.78
Grasslands	AVHRR	NASA NDVI	Grassland	0.39	0.34	0.56	0.41	0.57	0.64	0.60	0.56	0.38	0.60
Central Mixedwood	MODIS	NASA EVI	Mixed forest	0.42	0.30	0.56	0.65	0.61	0.58	0.51	0.56	0.60	0.69
Central Mixedwood	MODIS	NASA NDVI	Mixed forest	0.29	0.16	0.30	0.29	0.29	0.19	0.08	0.12	0.09	0.27
Central Mixedwood	MODIS	Pickell NDVI	Mixed forest	0.66	0.65	0.63	0.55	0.57	0.59	0.52	0.56	0.50	0.55
Central Mixedwood	AVHRR	NASA EVI	Mixed forest	0.42	0.34	0.69	0.68	0.76	0.78	0.67	0.68	0.62	0.52
Central Mixedwood	AVHRR	NASA NDVI	Mixed forest	0.35	0.40	0.29	0.29	0.27	0.27	0.20	0.20	0.20	-0.01
Montane	MODIS	NASA EVI	Conifer forest	0.07	0.03	0.15	0.15	0.14	0.24	0.23	0.17	0.15	0.15
Montane	MODIS	NASA NDVI	Conifer forest	0.74	0.61	0.62	0.65	0.61	0.59	0.39	0.54	0.41	0.57
Montane	MODIS	Pickell NDVI	Conifer forest	0.54	0.58	0.52	0.58	0.38	0.53	0.54	0.59	0.62	0.73
Montane	AVHRR	NASA EVI	Conifer forest	0.21	0.18	0.35	0.40	0.36	0.44	0.38	0.38	0.30	0.28
Montane	AVHRR	NASA NDVI	Conifer forest	-0.07	-0.15	0.08	0.15	0.17	0.16	0.09	0.14	0.11	0.05